

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 101781012
Source: IFW0
Date Processed by STIC: 2-4-05

ENTERED



IFWO

RAW SEQUENCE LISTING

DATE: 02/04/2005

PATENT APPLICATION: US/10/781,012

TIME: 16:12:10

Input Set : A:\50508-2200.txt

Output Set: N:\CRF4\02042005\J781012.raw

```

3 <110> APPLICANT: Emory University
5 <120> TITLE OF INVENTION: Mammalian Cell Lines Specifically Deficient in O-Linked
6   Glycosylation
8 <130> FILE REFERENCE: 050508-2200
C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/781,012
C--> 10 <141> CURRENT FILING DATE: 2004-02-18
10 <150> PRIOR APPLICATION NUMBER: 60/455,365
11 <151> PRIOR FILING DATE: 2003-03-17
13 <160> NUMBER OF SEQ ID NOS: 8
15 <170> SOFTWARE: PatentIn version 3.2
17 <210> SEQ ID NO: 1
18 <211> LENGTH: 1047
19 <212> TYPE: DNA
20 <213> ORGANISM: Homo Sapiens
22 <400> SEQUENCE: 1
23 atggcagaga aggtgctggt aacaggtggg gctggctaca ttggcagcca cacggtgctg      60
25 gagctgctgg aggctggcta cttgcctgtg gtcacgcata acttccataa tgccttccgt      120
27 ggagggggct ccctgcctga gagcctgcgg cgggtccagg agctgacagg ccgctctgtg      180
29 gagtttgagg agatggacat tttggaccag ggagccctac agcgtctctt caaaaagtac      240
31 agcttttatgg cggtcaccca ctttgcgggg ctcaaggccg tgggcgagtc ggtgcagaag      300
33 cctctggatt attacagagt taacctgacc gggaccatcc agcttctgga gatcatgaag      360
35 gccacggggg tgaagaacct ggtgttcagc agctcagcca ctgtgtacgg gaacccccag      420
37 tacctgcccc ttgatgaggc ccaccccacg ggtggttgta ccaaccctta cggcaagtcc      480
39 aagttcttca tcgaggaaat gatccgggac ctgtgccagg cagacaagac ttggaacgca      540
41 gtgctgctgc gctatttcaa cccacaggtg gcccatgcct ctggctgcat tggtagaggat      600
43 cccaggggca tacccaacaa cctcatgcct tatgtctccc aggtggcgat cgggcgacgg      660
45 gaggccctga atgtctttgg caatgactat gacacagagg atggcacagg tgtccgggat      720
47 tacatccatg tcgtggatct ggccaagggc cacattgcag ccttaaggaa gctgaaagaa      780
49 cagtgtggct gccggatcta caacctgggc acgggcacag gctattcagt gctgcagatg      840
51 gtccaggcta tggagaaggc ctctgggaag aagatcccgt acaagggtgt ggacggcg      900
53 gaaggtgatg tggcagccta ttacgccaac ccagcctgg cccaagagga gctggggtgg      960
55 acagcagcct tagggctgga caggatgtgt gaggatctct ggcgctggca gaagcagaat     1020
57 ccttcaggct ttggcacgca agcctga                                     1047
60 <210> SEQ ID NO: 2
61 <211> LENGTH: 348
62 <212> TYPE: PRT
63 <213> ORGANISM: Human polypeptide
65 <400> SEQUENCE: 2
67 Met Ala Glu Lys Val Leu Val Thr Gly Gly Ala Gly Tyr Ile Gly Ser
68 1          5          10          15
71 His Thr Val Leu Glu Leu Leu Glu Ala Gly Tyr Leu Pro Val Val Ile
72          20          25          30
75 Asp Asn Phe His Asn Ala Phe Arg Gly Gly Gly Ser Leu Pro Glu Ser

```

RAW SEQUENCE LISTING

DATE: 02/04/2005

PATENT APPLICATION: US/10/781,012

TIME: 16:12:10

Input Set : A:\50508-2200.txt

Output Set: N:\CRF4\02042005\J781012.raw

```

76          35          40          45
79 Leu Arg Arg Val Gln Glu Leu Thr Gly Arg Ser Val Glu Phe Glu Glu
80          50          55          60
83 Met Asp Ile Leu Asp Gln Gly Ala Leu Gln Arg Leu Phe Lys Lys Tyr
84 65          70          75          80
87 Ser Phe Met Ala Val Ile His Phe Ala Gly Leu Lys Ala Val Gly Glu
88          85          90          95
91 Ser Val Gln Lys Pro Leu Asp Tyr Tyr Arg Val Asn Leu Thr Gly Thr
92          100          105          110
95 Ile Gln Leu Leu Glu Ile Met Lys Ala His Gly Val Lys Asn Leu Val
96          115          120          125
99 Phe Ser Ser Ser Ala Thr Val Tyr Gly Asn Pro Gln Tyr Leu Pro Leu
100          130          135          140
103 Asp Glu Ala His Pro Thr Gly Gly Cys Thr Asn Pro Tyr Gly Lys Ser
104 145          150          155          160
107 Lys Phe Phe Ile Glu Glu Met Ile Arg Asp Leu Cys Gln Ala Asp Lys
108          165          170          175
111 Thr Trp Asn Ala Val Leu Leu Arg Tyr Phe Asn Pro Thr Gly Ala His
112          180          185          190
115 Ala Ser Gly Cys Ile Gly Glu Asp Pro Gln Gly Ile Pro Asn Asn Leu
116          195          200          205
119 Met Pro Tyr Val Ser Gln Val Ala Ile Gly Arg Arg Glu Ala Leu Asn
120          210          215          220
123 Val Phe Gly Asn Asp Tyr Asp Thr Glu Asp Gly Thr Gly Val Arg Asp
124 225          230          235          240
127 Tyr Ile His Val Val Asp Leu Ala Lys Gly His Ile Ala Ala Leu Arg
128          245          250          255
131 Lys Leu Lys Glu Gln Cys Gly Cys Arg Ile Tyr Asn Leu Gly Thr Gly
132          260          265          270
135 Thr Gly Tyr Ser Val Leu Gln Met Val Gln Ala Met Glu Lys Ala Ser
136          275          280          285
139 Gly Lys Lys Ile Pro Tyr Lys Val Val Ala Arg Arg Glu Gly Asp Val
140          290          295          300
143 Ala Ala Tyr Tyr Ala Asn Pro Ser Leu Ala Gln Glu Glu Leu Gly Trp
144 305          310          315          320
147 Thr Ala Ala Leu Gly Leu Asp Arg Met Cys Glu Asp Leu Trp Arg Trp
148          325          330          335
151 Gln Lys Gln Asn Pro Ser Gly Phe Gly Thr Gln Ala
152          340          345
155 <210> SEQ ID NO: 3
156 <211> LENGTH: 1017
157 <212> TYPE: DNA
158 <213> ORGANISM: Eschenichia coli
160 <400> SEQUENCE: 3
161 atgagagttc tggttaccgg tggtagcggt tacattggaa gtcatacctg tgtgcaatta      60
163 ctgcaaaacg gtcatgatgt catcattctt gataacctct gtaacagtaa gcgcagcgta      120
165 ctgcctgtta tcgagcgttt aggcggcaaa catccaacgt ttgttgaagg cgatatcgt      180
167 aacgaagcgt tgatgaccga gatcctgcac gatcacgcta tcgacaccgt gatccacttc      240
169 gccgggctga aagccgtggg cgaatcggtg caaaaaccgc tggaatatta cgacaacaat      300

```

RAW SEQUENCE LISTING

DATE: 02/04/2005

PATENT APPLICATION: US/10/781,012

TIME: 16:12:10

Input Set : A:\50508-2200.txt

Output Set: N:\CRF4\02042005\J781012.raw

```

171 gtcaacggca ctctgcgcct gattagcgcc atgcgcgccg ctaacgtcaa aaactttatt 360
173 tttagctcct ccgccaccgt ttatggcgat cagcccaaaa ttccatacgt tgaaagcttc 420
175 ccgaccggca caccgcaaag cccttacggc aaaagcaagc tgatggtgga acagatcctc 480
177 accgatctgc aaaaagccca gccggactgg agcattgccc tgctgcgcta cttcaacccg 540
179 gttggcgcgc atccgtcggg cgatatgggc gaagatccgc aaggcattcc gaataacctg 600
181 atgccatata tcgcccaggt tgctgtaggc cgtcgcgat cgctggcgat ttttggtaac 660
183 gattatccga ccgaagatgg tactggcgta cgcgattaca tccacgtaat ggatctggcg 720
185 gacgggtcacg tcgtggcgat ggaaaaactg gcgaacaagc caggcgtaaca catctacaac 780
187 ctcggcgctg gcgtaggcaa cagcgtgctg gacgtgggta atgccttcag caaagcctgc 840
189 ggcaaacccg ttaattatca ttttgcaccg cgtcgcgagg gcgaccttcc ggcctactgg 900
191 gcggacgcca gcaaaagccga ccgtgaactg aactggcgcg taacgcgcac actcgatgaa 960
193 atggcgcgagg acacctggca ctggcagtca cgccatccac agggatatcc cgattaa 1017
196 <210> SEQ ID NO: 4
197 <211> LENGTH: 338
198 <212> TYPE: PRT
199 <213> ORGANISM: Eschesichia Coli
201 <400> SEQUENCE: 4
203 Met Arg Val Leu Val Thr Gly Gly Ser Gly Tyr Ile Gly Ser His Thr
204 1 5 10 15
207 Cys Val Gln Leu Leu Gln Asn Gly His Asp Val Ile Ile Leu Asp Asn
208 20 25 30
211 Leu Cys Asn Ser Lys Arg Ser Val Leu Pro Val Ile Glu Arg Leu Gly
212 35 40 45
215 Gly Lys His Pro Thr Phe Val Glu Gly Asp Ile Arg Asn Glu Ala Leu
216 50 55 60
219 Met Thr Glu Ile Leu His Asp His Ala Ile Asp Thr Val Ile His Phe
220 65 70 75 80
223 Ala Gly Leu Lys Ala Val Gly Glu Ser Val Gln Lys Pro Leu Glu Tyr
224 85 90 95
227 Tyr Asp Asn Asn Val Asn Gly Thr Leu Arg Leu Ile Ser Ala Met Arg
228 100 105 110
231 Ala Ala Asn Val Lys Asn Phe Ile Phe Ser Ser Ser Ala Thr Val Tyr
232 115 120 125
235 Gly Asp Gln Pro Lys Ile Pro Tyr Val Glu Ser Phe Pro Thr Gly Thr
236 130 135 140
239 Pro Gln Ser Pro Tyr Gly Lys Ser Lys Leu Met Val Glu Gln Ile Leu
240 145 150 155 160
243 Thr Asp Leu Gln Lys Ala Gln Pro Asp Trp Ser Ile Ala Leu Leu Arg
244 165 170 175
247 Tyr Phe Asn Pro Val Gly Ala His Pro Ser Gly Asp Met Gly Glu Asp
248 180 185 190
251 Pro Gln Gly Ile Pro Asn Asn Leu Met Pro Tyr Ile Ala Gln Val Ala
252 195 200 205
255 Val Gly Arg Arg Asp Ser Leu Ala Ile Phe Gly Asn Asp Tyr Pro Thr
256 210 215 220
259 Glu Asp Gly Thr Gly Val Arg Asp Tyr Ile His Val Met Asp Leu Ala
260 225 230 235 240
263 Asp Gly His Val Val Ala Met Glu Lys Leu Ala Asn Lys Pro Gly Val
264 245 250 255

```

RAW SEQUENCE LISTING

DATE: 02/04/2005

PATENT APPLICATION: US/10/781,012

TIME: 16:12:10

Input Set : A:\50508-2200.txt

Output Set: N:\CRF4\02042005\J781012.raw

```

267 His Ile Tyr Asn Leu Gly Ala Gly Val Gly Asn Ser Val Leu Asp Val
268                260                265                270
271 Val Asn Ala Phe Ser Lys Ala Cys Gly Lys Pro Val Asn Tyr His Phe
272                275                280                285
275 Ala Pro Arg Arg Glu Gly Asp Leu Pro Ala Tyr Trp Ala Asp Ala Ser
276                290                295                300
279 Lys Ala Asp Arg Glu Leu Asn Trp Arg Val Thr Arg Thr Leu Asp Glu
280 305                310                315                320
283 Met Ala Gln Asp Thr Trp His Trp Gln Ser Arg His Pro Gln Gly Tyr
284                325                330                335
287 Pro Asp
291 <210> SEQ ID NO: 5
292 <211> LENGTH: 30
293 <212> TYPE: DNA
294 <213> ORGANISM: Human DNA
296 <400> SEQUENCE: 5
297 ggtgatgtgg cagcctatta cgccaacccc                               30
300 <210> SEQ ID NO: 6
301 <211> LENGTH: 33
302 <212> TYPE: DNA
303 <213> ORGANISM: Human DNA
305 <400> SEQUENCE: 6
306 gctggggttg gcgtaatagg ctgccacatc acc                               33
309 <210> SEQ ID NO: 7
310 <211> LENGTH: 40
311 <212> TYPE: DNA
312 <213> ORGANISM: Mouse DNA
314 <400> SEQUENCE: 7
315 cgcgacgcgt atccgtggcg ggacgaatgg acacagcaac                       40
318 <210> SEQ ID NO: 8
319 <211> LENGTH: 37
320 <212> TYPE: DNA
321 <213> ORGANISM: Mouse DNA
323 <400> SEQUENCE: 8
324 cgcgaagctt atcggtccg ctatgcgacg tgaggcc                               37

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/781,012

DATE: 02/04/2005

TIME: 16:12:11

Input Set : A:\50508-2200.txt

Output Set: N:\CRF4\02042005\J781012.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application No

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date